**Module-1:** **Analysis of a Betting Strategy in Sports**

*By: Nithin Reddy.P (NEU ID:002896440)*

*Subject:ALY6050: Introduction to Enterprise Analytics*

*Under the guidance of Shahram Sattar*

*Submission Date:25-Feb-2024*

# **Introduction:**

In this assignment, we have two teams Boston Red Sox and Yankees playing a series, in a home as well away games. The winner should win two out of three games or three out of five games. The main aim is to apply probability theory for a betting simulation.

# **Problem:**

Teams from NY and Boston are playing a three match series and the first team that wins two games will win the series. 0.6 is the probability of Red winning a game at home and for Yankees its 0.57 at their home turf. If a bet is placed you would win $500 if Red wins and loose $520 in the other case.

# **Facts:**

1. 0.57 is the probability of N.K.Y wining at home and for an away game it is 0.43
2. For Sox 0.6 is probability winning a home game and 0.4 for a touring game
3. Gains if Red Sox wins is $500
4. Loss if Red Sox lost is $520

# **Part-1:**

## Scenario-1: First game is played in Boston, second in New York and the last is at Boston.

### For B.R.S to Win the series:

It can be calculated by adding the probability of Red Sox fining the first two games, wins in first and last game as well as wins in second and last game. From Fig-1, we could see that the probability of wining the series is 0.5664 (sum of respective probabilities are 0.258, 0.2052 and 0.1032).

Probability distribution of Net win and calculation of expected net win and variance:

We are aware that $500 is the gain if Red Sox wins and a loose $520 is case of a series loss. We have identified the 0. 5664 as probability of wining the series, hence the probability of the loss would be 0.4336 (1- probability of wining the series). From Fig-2, we could see that the respective values and probabilities are assigned to vectors and expected net win is calculated using formula E(X)=∑i​xi​⋅P(X=xi​) I,e 57.728. The variance is calculated using formula Var(X)=∑i=1n​(xi​−μ)2⋅pi​ I,e 255512.9, thestandard deviation is calculated using formula Sqrt(Var(X)), ie 505.4829.

### Estimating Expected Net win using 95% confidence interval:

Please refer to Fig-3, a random sample is generated using rnorm function with the available mean and standard deviation. The sample mean, standard error, and z score for 95% interval have be calculated. By sung the sample mean, z score and standard error we have calculated the upper and lower bounds of the confidence interval have been found and we could see that the confidence interval (48.95895, 68.72811) contains the expected net win (57.728).

### Frequency Distribution and Chi-Square Goodness of fit:

Please refer to Fig-4, the generated random values haven been divided into bins and counting the number of frequencies in each bin. Later, the expected frequencies are calculated using cumulative distribution function. Finally, chi-square goodness of fit is calculated for distributions of observed and expected frequencies. The hypothesis is,

H0= There is no significant difference between the observed frequencies and the expected frequencies

H1= There is significant difference between the observed frequencies and the expected frequencies

Form the output, we could clearly see that the p-value is greater than significance level (0.05), hence we fail to reject the null hypothesis, concluding that observed frequencies are consistent with expected frequencies.

### Analysis of betting strategy of Scenario-1:

Based on the analysis the betting strategy is favourable and the below facts support the strategy, suggesting potential profitability in betting on Red Sox:

1. Probability of Red Sox winning the 3-match series is 0.5664
2. The obtained net win by a series win is $57.728 with a SD of 505.4829
3. The confidence interval of (48.95895, 68.72811) contains the expected net win (57.728).
4. The Chi-square good ness-of-fit suggest that observed frequencies closely match with expected frequencies, indicating a reliable strategy

# **Part-2:**

## Scenario-2: First game is played in New York, second in Boston and the last is at New York.

Please look at Figures-6,7, 8 and 9, the same steps have been performed like scenario-1 and please refer to the below findings.

### Analysis of betting strategy of Scenario-2:

Findings:

1. The probability of Red Sox wining is 0.47902
2. The obtained expected net win is $-31.3996 with standard deviation of $509.5508
3. The confidence interval (-50.1049, -30.06679) contains expected net win ($81.75)
4. The Chi-square good ness of fit suggest that observed frequencies closely match with expected frequencies, indicating a reliable strategy

The probability of Red Sox winning and expected net win changed from scenatrio-1 to scenario-2 but based on the above findings we can conclude that the betting strategy is not favourable, suggesting a loss of $-31.3.

# **Part-3:**

## Scenario-3:If a series is a best-of-five series the team that wins 3 games take the series, the games alter between Boston and NY, the first is at Boston. Please refer to Figures-10,-11, -12 and -13, the same steps have been performed similar to scenario-1 and kindly refer to the below findings.

1. The probability of Red Sox wining is 0.48174
2. The obtained expected net loss is $-28.6252 with standard deviation of $ 501.6774
3. The confidence interval (-32.98235, 13.34523) contains expected new win -28.6252
4. The Chi-square good ness of fit suggest that observed frequencies closely match with expected frequencies, indicating a reliable strategy

The probability of Red Sox winning is less than 50% and we would incur a loss of $-28.62, based on the observations we conclude that the betting strategy is not favourable as the outcome suggests a loss.

# **Conclusion:**

In conclusion, the betting strategy was examined under three scenarios. In the first scenario, the strategy was favourable. The probability of Red Sox winning as high, the expected net win as high and the chi-squared good ness of fit suggested the observed and expected frequencies closely match, which validates the strategy. On the other hand, the second and third scenario indicated that the strategy was not favourable as the net win as negative, indications a loss. Hence, it is crucial to consider and monitor the order of games and format of the series to formulate a better betting strategy.

**Appendix:**

## **Scenario-1**

### **Fig-1: Red Sox wins the series**

**A screenshot of a computer

Description automatically generated**

### Fig-2: Net win and Variance

**A screenshot of a computer code

Description automatically generated**

### Fig-3: Estimating net win using 95% confidence interval:

**A screenshot of a computer code

Description automatically generated**

### Fig-4: Frequency dist and Chi-Square Goodness of fit:

A screenshot of a computer program

Description automatically generated

## Scenario-2

### Fig-6: Probability of R.S winning the series

A screenshot of a computer code

Description automatically generated

### Fig-7: Expected Net win and Variance

A screenshot of a computer program

Description automatically generated

### Fig-8: Estimating Expected Net win using 95% confidence interval:

A screenshot of a computer code

Description automatically generated

### Fig-9: Frequency Distribution and Chi-Square Goodness of fit:

A screenshot of a computer program

Description automatically generated

## Scenario-3:

### Fig-10: Probability of Red Sox winning the series

A screenshot of a computer code

Description automatically generated

### Fig-11: Expected Net win and Variance

A screenshot of a computer program

Description automatically generated

### Fig-12: Estimating Expected Net win using 95% confidence interval:

A screenshot of a computer code

Description automatically generated

### Fig-13: Frequency Distribution and Chi-Square Goodness of fit:

A screenshot of a computer program

Description automatically generated

# **References:**

Probability course: Expected Value and Variance; [online]: <https://www.probabilitycourse.com/chapter4/4_1_2_expected_val_variance.php>

Open Stax: Expected Value and Standard Deviation; [online]: <https://openstax.org/books/statistics/pages/4-2-mean-or-expected-value-and-standard-deviation#:~:text=To%20find%20the%20expected%20value,%E2%88%91%20x%20P%20(%20x%20)%20>.